SIMDELTA -A MICROSIMULATION APPROACH TO HOUSEHOLD LOCATION MODELLING

Olga Feldman*[1], David Simmonds*, Vassilis Zachariadis*, Roger Mackett**, Mélanie Bosredon*, Emma Richmond*, James Nicoll

* David Simmonds Consultancy , ** University College London

ABSTRACT

The paper provides results on the household location modelling project which the UK Department for Transport commissioned as part of longer-term research into the factors underlying the demand for transport and the consequences of transport change. The new SimDELTA model design is based on replacing the household and individual components of the existing land-use modelling package DELTA with a system of microsimulation components. The treatment of transport, job location and housing supply are currently are exogenous to the model and based on previous modelling work. The proposed model has been calibrated and validated for the South and West Yorkshire, UK.

1. INTRODUCTION

Land-use and transport models are widely used as decision support systems in urban planning and management, as well as in public policy formation and transport forecasting. Proceeding from a simulation model, we can appraise the advantages of theories relating to urban phenomena and analyze the application of policy strategies (for example, growth management, congestion, and transport demand) to a variety of scenarios for urban prospects. Land-use and transport models are composed of independent land-use and travel models, with methods for integrating them into a single land-use and transport interaction model. Land-use models are applied to simulate demographic and economic transition in land-based activities, while transport models (in particular, travel demand models) are used to simulate travel patterns on a transport network.

The UK Department for Transport (DfT) has identified greater understanding of relationships between land use and transport as a key to forecasting and managing the demand for travel. This subject impacts of a number of policy areas but particularly that of tackling congestion. The paper provides results on the household location modelling project which the DfT commissioned as part of longer-term research into the factors underlying the demand for transport and the consequences of transport change.

[1] Corresponding author: Suite 14, Millers Yard, Mill Lane, Cambridge, CB2 1RQ, England, Tel: +44 1223 316098, Fax: +44 1223 313893, E-mail: olga@davidsimmonds.com,

This project is not intended to deliver a full land-use/transport interaction model, i.e. it is not expected to simulate the feedback mechanism from the land-use model to transport model, but only to model household location and commuting that responds to transport. The proposed model SimDELTA is mainly intended to forecast the future number of households living in each zone of the study area with adequate elements of household and household member characteristics and possibilities to predict household population and car ownership. It also considers all the processes which change household composition in each zone over time, from the simple, such as ageing, to the complex, such as couple formation and dissolution.

The model also takes into account the transport system and its impact on accessibility and therefore on households, changes in the supply and characteristics of housing and changes in the number and distribution of workplaces. Households and their members go through a series of processes associated with the long-term response to changes in transport and related processes.

The SimDELTA model design is based on replacing the household and individual components of the existing land-use modelling package DELTA, which has been developed at David Simmonds Consultancy Ltd., with a system of microsimulation components partly based on the microsimulation model MASTER (Micro-Analytical Simulation of Transport, Employment and Residence) model, which was previously developed at the University College London (Mackett, 1990, 1992, 1993).

The proposed model has been calibrated and validated for the South and West Yorkshire, UK. The reason for choosing this area is the existence of the modelling system for this area known as SWYSM (South and West Yorkshire Strategic Model). SWYSM is a land-use/transport interaction model based on the DELTA land-use/economic model software (devised and maintained by the David Simmonds Consultancy) and the START strategic transport model (developed and maintained by MVA Consultancy). SWYSM was one of two models developed for the South and West Yorkshire Multi-Modal Study. The other, known as the Detailed Transport Model (DTM), provided demand and supply data for SWYSM and was used for more detailed but more conventional analysis of transport interventions. Further details of the SWYSM model can be found in MVA (2001). Both SWYSM models have been used in the present project to get the transport inputs to DELTA. In the future the SimDELTA model can be linked directly to the SWYSM model to have the full land-use and transport interaction model.

Microsimulation techniques have been used in this study to simulate the way the population changes through time. Microsimulation implicates modelling the choice processes and changes of a large sample of individual households and household members, considering the members of the sample one at a time and predicting a specific outcome for each choice or change for each member. This contrasts with conventional "aggregate" modelling, which considers all the households in a particular category, and predicts what proportion of them will make each choice.

The following processes are presented in the microsimulation model.

- Individual demographics and other changes: ageing, survival, birth/multiple birth, entering labour market, re-entering labour market, retiring from labour market, educational status, acquiring driving licence, becoming permanently sick, moving to institution.
- Household changes: separation, couple formation, marriage, absence from households, student households and other shares, obtaining/losing car, household income, housing affordability.
- Household location: seeking to move, housing tenure choice, dwelling choice, housing prices or rents, location choice, household location/relocation, household in/out migration.
- Employment: job supply, seeking to change job, job and workplace choices, wages, accepting/rejecting candidate, accepting/rejecting job.

In subsequent sections the following aspects of the model are discussed. Chapter 2 describes the modelling area. Chapter 3 give a brief introduction into the static spatial modelling methodology which was used to create the initial micro dataset. Chapter 4 specifies the overall structure of the dynamic model and the design of the microsimulation components whist Chapter 5 contains some examples of the results. Finally, Chapter 6 concludes.

2. MODELLING AREA

The SimDELTA study area is part of the South and West Yorkshire, UK (see Figure 1). The definition of the case study area within the South and West Yorkshire takes account of the work on the Functional Areas and Regions which MVA Consultancy and David Simmonds Consultancy Ltd carried out for the UK Department for Transport (MVA and DSC, 2005). The Functional Area analysis provided a catalogue of possible zoning systems at different levels of aggregation. At any chosen level of aggregation, the set of zones is based on identifying areas of relatively high self-containment in the travel-to-work patterns. As it can be seen from Figure 1, the modelling area boundaries do not always coincide with the local authority boundaries.

The functional areas were defined using 2001 data. The level of aggregation was 1000 Functional Areas for England and Wales. For the calibration purpose the model was run from 1991 to 2001 and it was necessary to model the set of 1991 wards which correspond to those 2001-defined areas. This set was defined as the two hundreds and eighty one 1991 wards whose centroid fell within the boundary of the 2001 study wards, plus an additional four 1991 wards which were included in this set because they had significant overlap with the 2001 study zone. The 1991-based model is therefore running on 283 fully modelled zones, each one being a single 1991 ward.

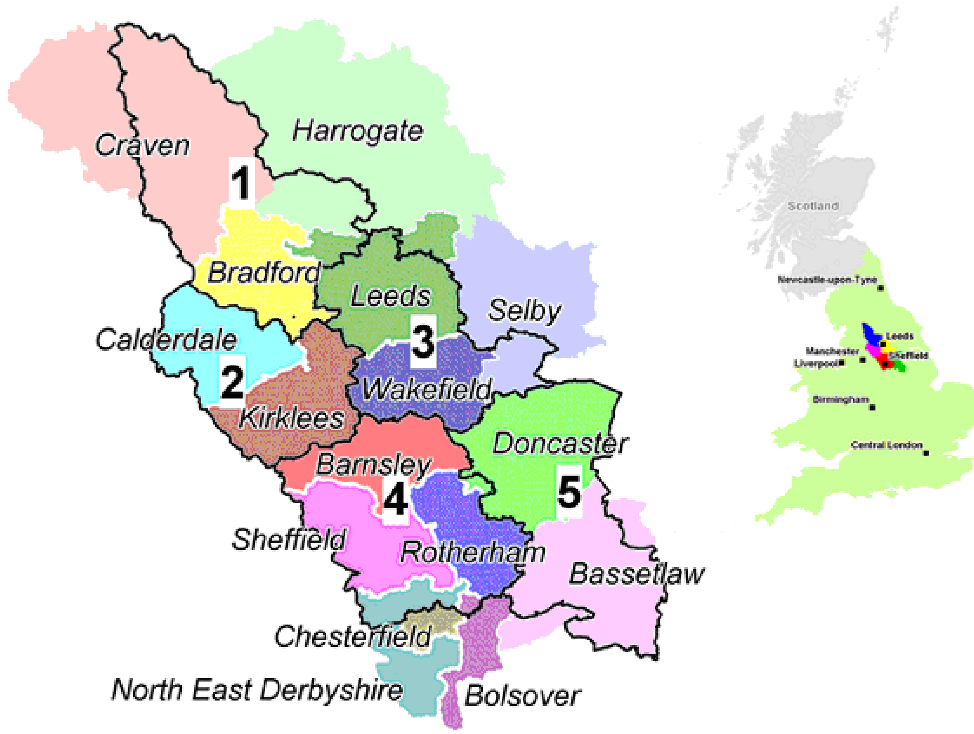

Figure 1 Modelling Area.

There are 18 external zones in the microsimulation part of SimDELTA which represent the rest of the South and West Yorkshire and also Greater Manchester, Humberside, Lincoln, Nottingam, Derby, Stoke, East Lancashire and Yorkshire. We do not model population at the microsimulation level in these areas but we do model population in these at the aggregate level.

3. THE STATIC SPATIAL MICROSIMULATION MODELLING METHODOLOGY

One clear requirement of the microsimulation approach is to simulate the initial population. The UK Census of Population has been and remains the most reliable social accounting of people and housing in Britain and is an exceptional source of data for the social sciences (Dale, 1993; Rees et al, 2002). However, the restrictions on the release of Census data mean that it is not possible to obtain such a database directly, and therefore a synthetic database has to be created.

In Britain, the UK Census offices release Samples of Anonymised Records (SARs) from the Census of Population (Dale et al., 2000). The 1991 UK SARS is a 2% sample of individuals in households and communal establishments comprising 1.1 million records that contain information on all the topics asked in the census and limited information about other members of the household, and a 1% sample of households and individuals in those households comprising 215,000 households and the 542,000 individuals enumerated in them. For the purpose of this study we use a 1% sample of households and individuals in those households extracted from the 1991 Census of Population. Unfortunately, the 2001 Household Sample of Anonymised Records was not yet released when the project was carried out. The 2001 SARs also contains less details then 1991 SARs.

Table 1. Population in 1991 Census and 1991 Synthetic Databases for the Study Area

Local authority (whole or part)	**Population Static Model 1991**	**Population table sas035 Census 1991**	**Deviation from Census, %**	**Households Static model 1991 = sas27**
Barnsley	218,731	218,556	0.080	**87,087**
Bassetlaw	98,881	98,779	0.103	**39,218**
Bolsover	1,917	1,912	0.262	**778**
Bradford	451,308	451,050	0.057	**174,071**
Calderdale	189,520	189,444	0.040	**77,428**
Craven	36,615	36,554	0.167	**15,290**
Doncaster	285,264	285,022	0.085	**112,725**
Harrogate	4,344	4,347	-0.069	**1,638**
Kirklees	369,696	369,534	0.044	**146,864**
Leeds	648,841	648,572	0.041	**271,518**
North East Derbyshire	46,398	46,397	0.002	**18,027**
Rotherham	249,795	249,692	0.041	**97,851**
Selby	10,334	10,317	0.165	**3,857**
Sheffield	496,447	496,215	0.047	**210,925**
Wakefield	307,936	307,765	0.056	**123,465**
Total	**3,416,027**	**3,414,156**	**0.055**	**1,380,742**

To create a synthetic initial dataset we used a combinatorial optimisation method called simulated annealing (Metropolis et al., 1953). The simulated annealing modelling code in Java (Ballas, 2001) was used in order to estimate spatially disaggregated microdata for the study region electoral wards, using data from the 1991 Household Sample of Anonymised Records and Small Area Statistics (SAS) tables from the 1991 UK Census of Population. The simulated annealing procedure in a population spatial microsimulation context works by reading in SAS tables and SARs data, selecting at random households from the records of the SARs sub-set for Yorkshire and the Humber and East Midlands

Regions (selected with repetition) to population the SAS tables, applying a simulated annealing algorithm to find the best fitting set of households and saving the set of SARs households that best fits the SAS tables. In particular, the algorithm works by selecting an initial random sample of records until sufficient households are represented (see Feldman et al, 2005).

The base year in the model is 1991. The model was then run to 2001 and calibrated so that as to match the 2001 Census data as close as possible. The initial synthetic databases were created for 1991 and Table 1 compares tabulations of the aggregate population microdata with the data from the Census table SAS035. The comparison is simply in terms of the total population by local authority area (or part local authority area, where applicable). The total numbers of households in each ward within the modelling area in 1991 were taken from 1991 Census Table SAS 27, which means that the error between the static model output and the Census SAS 27 table is zero for each ward.

As can be seen from Table 1, the total population of the modelling area in the base year is 3,416,027 people or 1,380,742 households.

4. DYNAMIC MODEL IMPLEMENTATION

The output of the static microsimulation model provides the input for the dynamic microsimulation model. The dynamic model is based on DELTA and MASTER models. The summary of the DELTA package, which is based upon sub-models representing different processes of change, with most of the interactions between these processes being gradual rather than simultaneous, is described in Simmonds and Feldman (2005). The DELTA package and its existing application to South and West Yorkshire (Simmonds and Skinner, 2004) are used to provide aggregate modelling of non-household elements (e.g. employment and housing change), and the existing transport models from the same study are used to provide inputs to accessibility calculations.

The overall structure of the proposed model is presented in Figure 2, which shows the sequence in which the different DELTA components are applied. The logical linkages between the different components are much more complex, with many time-lagged relationships and, of course, many feedbacks both positive and negative.

We will briefly describe the processes of change that are represented in the model.

Individual demographics and other changes

- **Ageing.** The aging process is straightforward. The age of each person is increased by 1 each year

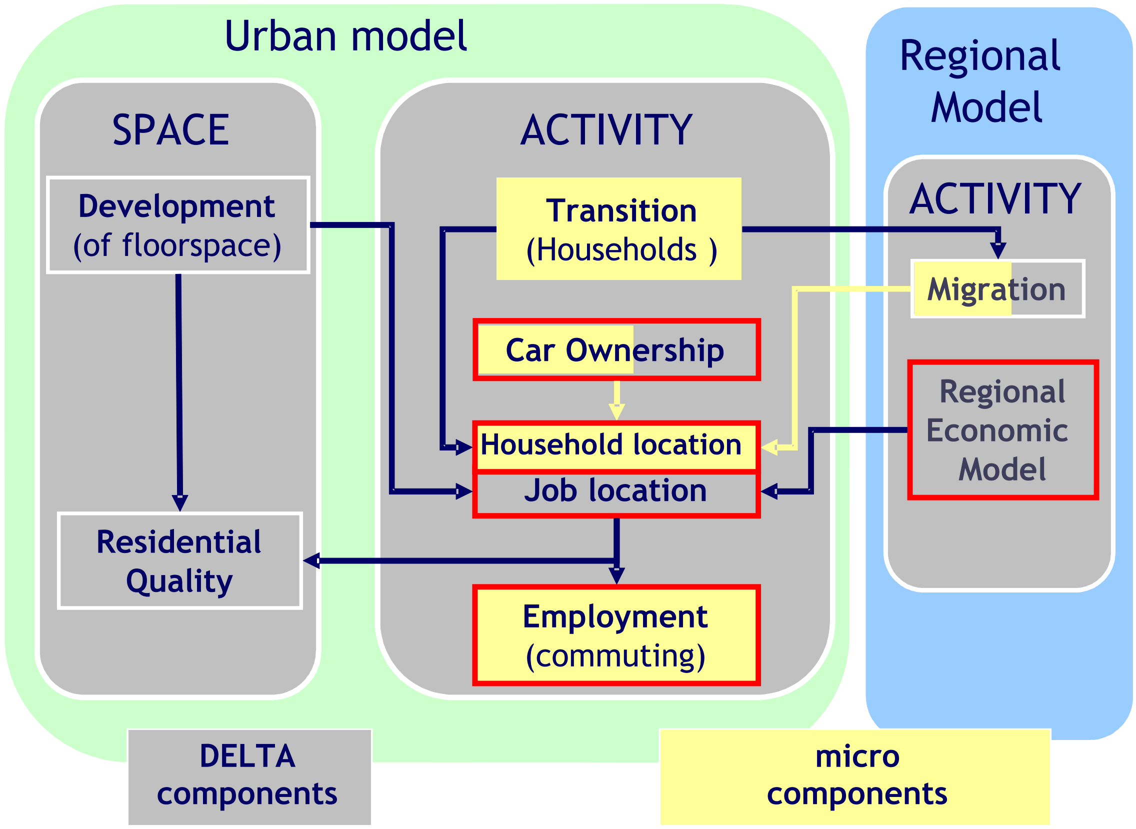

Figure 2 SimDELTA model structure within one time period.

- **Survival.** Monte Carlo simulation is used here, we assume that the probability of an individual surviving within the simulation period is a function of age and gender.

- **Birth/multiple birth.** Monte Carlo simulation is used here. Births are modelled using birth rates by five-year age group, ethnicity, and couple status. Then Monte Carlo simulation is used (with a constant probability) to see if twins are born (the possibility of triplets or more is ignored). The gender of the child is determined by Monte Carlo simulation with fixed probabilities for the ratio of males to females.

  The new individual's attributes are set as follows: age is zero, sex is determined probabilistically, couple status is single, ethnicity and location is that of the mother, all the other personal characteristics are undefined. In the next simulation period, the new individual is simulated along with the other individuals in the household.

- **Socio-economic status.** All adult persons within the model were given values representing their Socio-Economic Group (4 different groups) based on their SARs socio-economic group where applicable or using Monte Carlo simulation if their SARs socio-economic group was "not applicable" or "not adequately described”. Every child is automatically attributed the socio-economic group 0 - non-defined until reaching working age (16 years). On completing education, the person’s socio-economic group is updated accordingly. We use Monte Carlo simulation in order to allow people to move to other socio-economic groups throughout the rest of the working life. The probabilities in this process depend on age (5 age groups), sex and their current socio-economic group. Persons search for employment compatible with their socio-economic groups; if this is not available they expand their search accordingly. Applicant acceptance is also conditional by the job/worker socio-economic group match.

- **Educational status:** Education attributes extracted by UK census data are based on “level of highest qualification” values for persons over 16 years old and any person under 16 has no relevant education attribute. Information for newly processed persons is accordingly updated in the model based on his/her education status (student to 18, student to 21, etc). In order to be consistent with the national statistics, the model allows persons to become students at older ages.

- **Entering/re-entering labour market:** Each person in the model is assigned one of 11 possible economic statuses (not applicable, employee full time, employee part time, self-employed with employees, self-employed without employees, government scheme, unemployed, student, permanently sick, retired, other inactive). Once the person is employed, he or she normally stays employed but can change job, become retired, become permanently sick, become redundant or become “other inactive”. To address deliberate withdrawal from market due to family related matters (parenting role for mothers) any mother may “decide” to leave the market after the birth of a child. Decisions are made through a simple Monte-Carlo simulation. Successful departures from the market result in change in mother’s economic status from an economic active status to “Other Inactive”.

- **Re-entering labour market:** Persons whose economic status is inactive (but not students) have a fixed probability of re-entering the labour market. For females which gave birth to children, re-entry to the labour market is modelled through “positive” decisions of female persons of economic group “Other inactive” to become economic active. Probability of re-entry is based on their socio-economic group and age of their youngest child.

- **Redundancy:** Within an apparently static situation, jobs are usually being lost due to the decline/closure of individual firms/establishments, whilst an equal number of similar jobs is being created due to the growth of other firms/establishments in the same zone and sector. To represent this we use a "normal redundancy rate" (which varies across the four socio-economic groups). This rate is defined as the probability that a worker in that category will lose his/her job when there is no net change in the sector in the zone.

  The redundancy rates are applied to each worker by Monte Carlo Simulation. If a worker is made redundant, he/she will not be able to seek another job in the same year (unless he/she changes household and the new household is considered later that year). Newly redundant people are treated like previously unemployed. No natural churn process is implemented.

- **Retiring from labour market:** Monte Carlo simulation is used to determine whether the person is going to retire or not in a particular year as a function of age and gender. Most men retire at age 65 and women at 60. At age 75 all those people who were still employed, retire. For each worker who retires the model increases by one the corresponding number of vacancies. In principle, some people already retired may choose to come out of retirement if the financial incentive or necessity is sufficient, but we are not considering such movements.

- **Acquiring/losing driving licence:** Monte Carlo simulation is used here. Each person over 17 years old can acquire a driving licence if he/she did not have one or they can lose it if they had one. The probabilities in each case are driven by age and sex of people.

- **Becoming permanently sick:** Any economic active person may at any time become permanently sick. In this case the person leaves his job and the job market permanently. Probability to become permanently sick is based on age, sex and socio-economic group of persons.

- **Moving to institution.** We adjusted the simulated microdata for elderly population movements to institutions of care such as care homes, nursing homes etc. Higher mortality rates are applied to those residing in communal establishments. This assumption is based on the reasoning that many residents move to communal establishments when they can no longer look after themselves and tend to observe relatively short periods of stay. It is also assumed that most residents do not leave communal establishments to return to the household population (if they are going to return to regular household population they do so in the first year at a rate of 10%).

Household changes

- **Couple formation and marriage:** A male-dominant model is implemented when the event 'couple formation' is simulated for the man, using Monte Carlo simulation, as a function of age and couple status. The model does not account for same sex couples. A new couple can be married or cohabiting. There is a further process whereby cohabiting couples may marry. We assume that partners are generally found within the area of residence rather than throughout the full system. Nevertheless, there is a low probability that partners are found in different locations, and the model allows for migration for people who do form couples with persons from other regions. In this case one of the partners is moving to the location where the second partner is currently living. Monte Carlo simulation is used to identify who is moving.

- **Separation:** The separation of both married and cohabiting couples is modelled as a function of age and couple status using Monte Carlo simulation. Separation of married couples implies divorce.

- **Absence from households:** Monte Carlo simulation is used here. A young person can leave parent home and then to come back in a number of years.

- **Student only and shared households:** The following assumptions have been made: all student households are sharers; a household is a student household if and only if all occupants are students - regardless of their marital status. A household is a sharer household only if each occupant is either a student or is aged between 16 and 34 inclusive and single (not married or cohabiting) and has no children. Moreover a household is a sharer household only if there are more than one families living in it or if the (single) head of household marks it as such (person living alone in a household but seeking room-mates).
  We have introduced a process of checking whether a household is a student or a sharer household before running the processes which are sensitive to the household status (e.g. marriage process of new family's location depends on the status of marrying persons in old households). The following processes are affected by the status: single person leaving home, division, marriage, income simulation and employment change.

  The sharer definition within the static model (the synthetic population in the base year) was modified to include single person households that dwell in households with at least three empty rooms. This was done in order to model the sharer mobility process. The model is able to handle households that consist of more than one family and offers opportunities for household divisions. Divisions are restricted to couples leaving households where none of the members of the couple is head of household.

- **Obtaining/losing car:** The SWYSM land use model has a car ownership sub-model which works entirely in terms of the zonal probabilities of a household of a particular type owning no car, one car or tho-plus cars. These are effectively conditional probabilities - they imply that if such a household locates in a given zone, then it will have those probabilities of owing that many cars. These probabilities are passed to the microsimulation model and they are used to test whether a household will obtain or lose a car using Monte Carlo simulation.

- **Household income:** The household income is based on the gross individual incomes of the persons in the household and it is simulated as the sum of the incomes of each member of the household. Working persons are contributing their wages, while unemployed and retirees and permanent sick are contributing fixed sums based on their characteristics. Children are also adding a sum to account for children benefits and tax policies. Mature students involved in further education are considered to retain the wage of their last job.

- **Housing affordability:** Households are programmed to form budgets for buying or renting dwelling based on their tenure preferences, the values of the housing markets and the characteristics of available housing stock. In case of renting, household budgets are formed as a fixed proportion of the household gross income. Based on Monte-Carlo simulation each household is fixing a budget from 25 to 35 percent of the annual gross income for housing (renting).

  While renting budgets are based on gross household incomes, buying budgets are based on a number of parameters such as savings, net household income, outstanding mortgages from previous acquisitions and previous type of tenure. Savings are calculated each year for every household after subtracting costs of living (transport costs, foods and goods costs, taxing costs etc.) and housing costs (rent or mortgage) from annual housing income. Outstanding mortgages are passed on each year after subtracting a sum equal to a set proportion of the household annual income. The buying budget is formed by adding savings and – in case of owning a one – current dwelling's value after subtracting outstanding mortgages. Households are also adding to their budgets a new mortgage based on their current net income and the current multiplying factor. However, in case of outstanding mortgage the value of the mortgage is decreased by an equal amount. Moreover, students, unemployed and other various non-earner households are not eligible to get a new mortgage, while sharer and student households cannot change their tenure preference to own-occupiers.

Employment

- **Job supply**: The job supply for the people living in the microsimulation Study Area and working in the microsimulation Study Area or in the microsimulation external zones are output of the SWYSM model. A special interface was written in order to do it. Full-time and part-time jobs are presented in SimDELTA.

  Instead of retaining the tactic of handling employment implicitly and representing available and occupied jobs as figures for each area, we opted for a consistent to the rest of the model strategy and represent jobs as objects: each job contains attributes for current occupier, socioeconomic group of the occupation, current wage, location etc. The advantages of this solution are many: accountability, more information and most important expansibility to modelling or integrating to other land uses modelling (retailing, services etc.).

- **Main earner:** We have introduced the concept of the main earner, or earners, in a household which could have a strong influence on household decision of whether/where to relocate. The main earner is the person who is likely to be the main income-earner of the household - although he or she may be unemployed at a particular point in time. The new definition is intended to avoid problems with the inconsistency of "head of households" and "household reference person" in the 1991 and 2001 UK Censuses. We will assume that in any household with just one economically active person that person is the main earner. If a households contains a couple (married or cohabiting) who are both economically active, of equal socio-economic group or occupation, and in white-collar or skilled manual levels, then we will assume this to be a dual-career household, and we will assume that the members of the couple are joint main earners.

In all other households containing more than one economically active person, we will take the oldest person in or seeking full-time employment as the main earner. If no one is in or seeking full-time employment, we will simply take the oldest person in or seeking employment.

- **Job and workplace choices:** Unemployed, agents entering the labour market for the first time and those returning to it after a break (maternity leave) are looking for a job searching through areas after sorting them out by utility based on two parameters: generalised cost to travel form current residence and number of available jobs in area with specified target socioeconomic group attribute.

  Each agent is "applying" for a specified number of jobs in each area and searches only in a number of areas. Unemployed persons unsuccessful in finding a workplace that suits their preferences, are likely (based on Monte-Carlo simulation) to decide to search for jobs of other socioeconomic groups or other economic types (part-time instead of full-time and vice versa). In case an application is accepted the unemployed agent halts job-seeking and accepts the job.

- **Seek-to-change-job process**: The process followed by an agent that already occupies a job is similar to the one for unemployed and entering to the labour market agents. However, there are two major differences: in this case areas are sorted based on the generalised cost to travel from the area of their residence, number of jobs available in each area and the distance from current work location. Secondly, after a job application has been accepted, the employed agent is comparing the proposed wage to the one he/she receives from his/her current job. If the wage is lower the new job is rejected. This way it is ensured that agents only get jobs of salaries equal or higher than those earned by their old job.

- **Wages:** Wages are allocated to all working persons within the model. Each working person is now connected to a specified annual wage based on its characteristics and the job it occupies. The offered wages are based on agent's age, socioeconomic group and gender and a degree of randomness that distorts the offer from 80 to 120 per cent.

- **Accept/reject candidate, accept/reject job:** Monte Carlo simulation is used in both cases to decide whether the job was offered to the potential worker and whether it was accepted. The chances (again based on Monte-Carlo) of their applications been accepted are changing based on how much their profile matches the attributes of the job he or she applied for.

Household location

- **Housing stock:** Regarding modelling of the destination of relocations of households we consider the supply side availability of dwellings. The change in the number of dwellings is an exogenous input to SimDELTA model. For the years beyond 2001 the planning policy files have to be set up with the numbers of dwelling built or demolished. For forecasting beyond 2001, the default is that all dwellings continue to be of the same tenure as in the initial database or (if they are built after 2001) stays of the same tenure as they are initially assigned.

- Four dwelling types are modelled: detached house, semi-detached house, terraced house and flat. In the synthetic database each dwelling belongs to one of these four types (from SARS 1991). The dwellings also differ by the dwelling sizes (number of rooms). Currently we do not have information on the number of new dwellings built by dwelling type and the number of rooms therefore a "cloning" process is used to set these parameters for new dwellings: a random dwelling from the current database is chosen and the characteristics of this dwelling are copied to the new dwelling.

- **Seek to move:** After budget formation, households search within their preferred areas for a vacant house matching their preferences (tenure and budget) and sized within their size tolerance (usually one room tolerance). In order to avoid futile searches, households seeking relocation check whether their budget is over the expected minimum budget for a house of the required size in an area before searching it. Households search a fixed number of appropriate areas and if they are unsuccessful in renting a property, look for alternative tenure types before giving up. In the case a household is unsuccessful in relocating, it returns to its previous dwelling. Unsuccessful external in-migrants are destroyed.

  If a household finds a suitable vacant dwelling in an appropriate area within its budget it marks it as potential target. In the case that it finds another or multiple dwellings in the same area and again under its budget, the household always prefers the one closer to its budget, thus the most expensive one. This is done in order to maximise utility. Areas are sorted based on their utility. Utility is calculated based on area deprivation, distance for current area of residence, general accessibility of target area and generalised cost of target area to main-earners work place.

- **Housing tenure choice:** Household's choice of tenure is influenced by the supply of dwellings of each tenure type (this is fixed). Households unable to find accommodation of their preferred tenure within their budget constraints may in some cases switch to a different tenure type.

- **Dwelling choice:** The actual dwelling that is chosen by a moving household or individual must fit the required characteristics – the dwelling size. Households moving due to high room stress (too many people per room) can only move to a more suitable (lower room stress) property. If none are available they have to wait. If a suitable set of houses exist then the actual property chosen is randomly assigned.

- **Housing prices or rents:** The household location model requires "asking prices" to be set for owner-occupied dwellings which are being sold, and "asking rents" for dwellings which are being let. The rents are modelled as fractions of the sale price, ie to assume a fixed yield, plus market adjustments. The sale price is calculated using a hedonic price model based on
    - the price of a typical dwelling of a particular type in this zone (this data is available up to 2005, for future years the 2005 price is used and the inflation index is applied),
    - the location constant of this zone at in this year,
    - the cumulative from the base year inflation rate of housing prices,
    - the average price if a room in a dwelling of each type in the base year,
    - the size difference (the number of rooms between the present dwelling and the typical one) for each dwelling type,
    - the market change indicator which reflects the demand and supply sides.

  Rather than attempting to devise distinct models for rents, the plan is to set rents at a fixed proportion of the corresponding sale price, ie to assume a fixed yield, plus market adjustments.

- **Location choice:** A price or rent-based location model is implemented. Households trade off desirable housing, location and accessibility characteristics against price or rent, and price or rent are adjusted over time in response to changes in the balance between supply and demand.

- **Household location/relocation and migration:** The overall model sequence is presented in Figure 3. Job choice for main earners may occur before residential relocation (i.e. change of job by the main earner can lead to household relocation); job choice for others (and possibly for main earners) is considered after household relocation (in the next year).

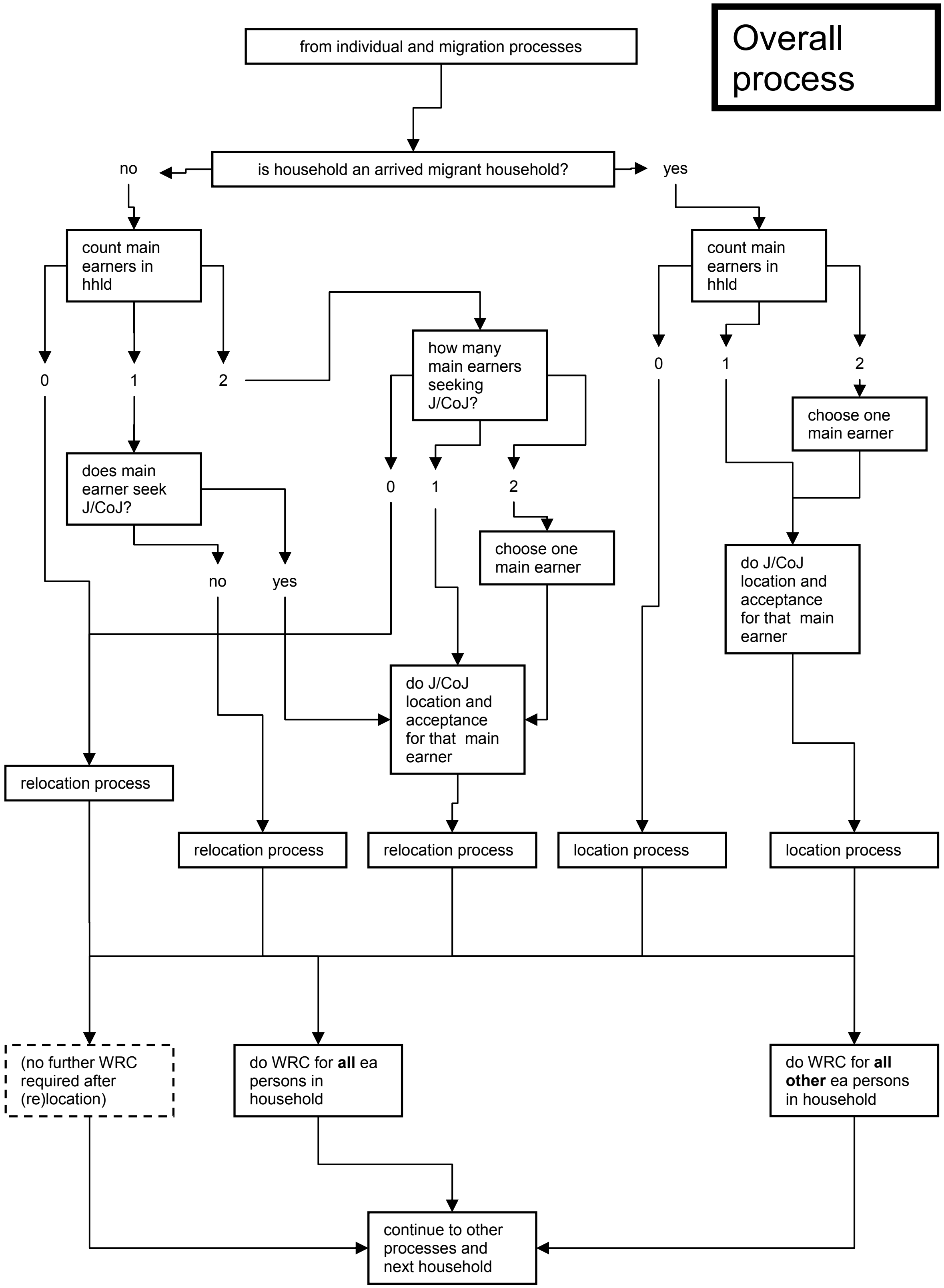

Figure 3: Overall process of residential/work choices

Before considering either change of job or change of dwelling, we test whether the household is going to migrate (make a longer-distance relocation necessarily involving a change of job(s)). If so, then they disappear from their existing area – and may reappear elsewhere in the model as migrants into another area.

The model also predicts change of socio-economic group for each economically active individual, as a function of qualification, age and being in employment (as a proxy for qualifications and experience) before considering change of job/dwelling. The model also predicts whether an economically active person is seeking a full-time or part-time job (including the possibility that an already-working person wishes to switch from one to the other).

For households which have not migrated, we consider possible job changes by main earners (which could have a strong influence on household decision of whether/where to relocate), household relocation, and possible job changes by other household members.

For households with one main earner, we first test whether he/she is seeking a new job (ie whether he/she is employed and seeking to change job, or is currently unemployed). If yes then we model work-related choices for that person and then we model relocation choices for the household, which will be influenced by the job location. Then (whether or not relocation results) we model work-related choices for any other workers or unemployed persons in the household. If the main earner is not seeking a new job, then we model location choices. If relocation occurs, we model job choices for all working members of the household; if relocation does not occur then we model whether-to-change job for the other working members of the household (if any other than main earner exist).

For households with two main earners, i.e. dual career households, we make some allowance for each main earner being restrained by the other.

For in-migrant households (whether in-migrants from another part of the microsimulation area, or from the rest of the world), the model considers, in turn, job seeking by the main earner(s), household location, and possible job seeking by other household members. This is an unconditional sequence - the main earner(s) must seek work before household location occurs (though they may fail to find work, in which case the household location model must work without any reference to workplace.)

5. INTERFACES BETWEEN THE AGGREGATE AND MICROSIMULATION MODELS

DELTA-to-Microsimulation interfaces

The present project set out to develop microsimulation models of household and individual choices which would replace the equivalent aggregate models within the

SWYSM DELTA model. At one stage the intention was literally to remove the household and individual choices from the operations of the SWYSM DELTA application. However, we subsequently identified the need to input to the microsimulation, the number of jobs by zone available to microsimulation residents. This left a need for the DELTA model to consider the competition for jobs between residents of the microsimulation area and residents of the rest of the region. Since this competition is already modelled in SWYSM DELTA, the obvious choice was to retain the whole of that model to produce the required job numbers, but with the DELTA household/person forecasts for the microsimulated area being overwritten with the microsimulation forecasts for each year of each model run.

An additional option has been programmed in DELTA which specifies that DELTA is to be run as part of SimDELTA. This controls

- reading additional definitions to match SWYSM zones to microsimulation zones;
- outputs of microsimulation data from standard DELTA programs;
- running additional interface programs where required;
- stopping at the appropriate point in the DELTA sub-model sequence each year for the microsimulation to be run; and when it has run,
- overwriting the standard DELTA outputs with the aggregated microsimulation results.

The interfaces from the aggregate model to the microsimulation model were set up for :

- migration forecasts,
- job forecasts,
- car ownership forecasts, and
- housing forecasts.

Microsimulation-to-DELTA interface

The microsimulation outputs first need to be aggregated to strategic zones and to DELTA classifications. This is done by the microsimulation model. A new DELTA program has been written to read in these microsimulation results and to superimpose them on the standard DELTA results for

- households and population in the microsimulated area,
- travel-to-work flows from the microsimulated area, whether to other parts of the microsimulated area or to zones outside that area.

SOME EXAMPLES OF THE RESULTS

The model provides a lot of very detailed outputs for population and households and the user can aggregate the output at any desired level. It is also possible to trace individuals, households, jobs and dwelling in time. For example, to obtain such information as who lived in a particular dwelling or who occupied a particular job in a particular year. In this paper we would like to show the model output on the average population change between

1991 and 2001 for the model Study area and some outputs for a test where extra dwellings were built in a particular zone as an example of model's capabilities.

Figure 4 shows the base population in 1991 and the results of a model run on the population change in 2001. The model was calibrated so as to match the 2001 UK Census data as best as possible. It is well known that microsimulation models produce different results if run with different seeds which produce different random numbers. It is advisable to do a number of runs and then to calculate average results on aggregate variables. In our model at the zone aggregate level for the total population the differences between two different tests are very small. Obviously, the results on the individual level can be very different.

As a demonstration of the model response to the supplementary housing development, a test was run where an additional 1000 new dwellings were introduced in to ward 05CCFX (Wombell North, Barnsley) in 1993. Such a test allows us to see whether current residents are likely to benefit from a new development or whether such a development has the effect of attracting new people to an area.

The model was run on a 10% base population sample and the results presented here were obtained by multiplying the model output by 10.

Table 2 shows the increase in the number of dwellings, the number of households and the population over the course of the model. There is an obvious increase in the number of dwellings from 1992 to 1993 combined with a jump in population and household numbers. The increase in population and number of households then continues until 1998 as the new dwellings which were initially vacant become occupied.

Table 2. Population change in ward 05CCFX from 1992 to 2001 as the result of the model test on the extra development in this zone.

	1992	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**
Dwellings	4500	5510	5520	5530	5540	5550	5560	5570	5580	5590
Population	10540	11600	11900	12460	12970	13380	14180	13740	13730	13780
Households	4180	4560	4680	4970	5180	5370	5550	5560	5570	5580

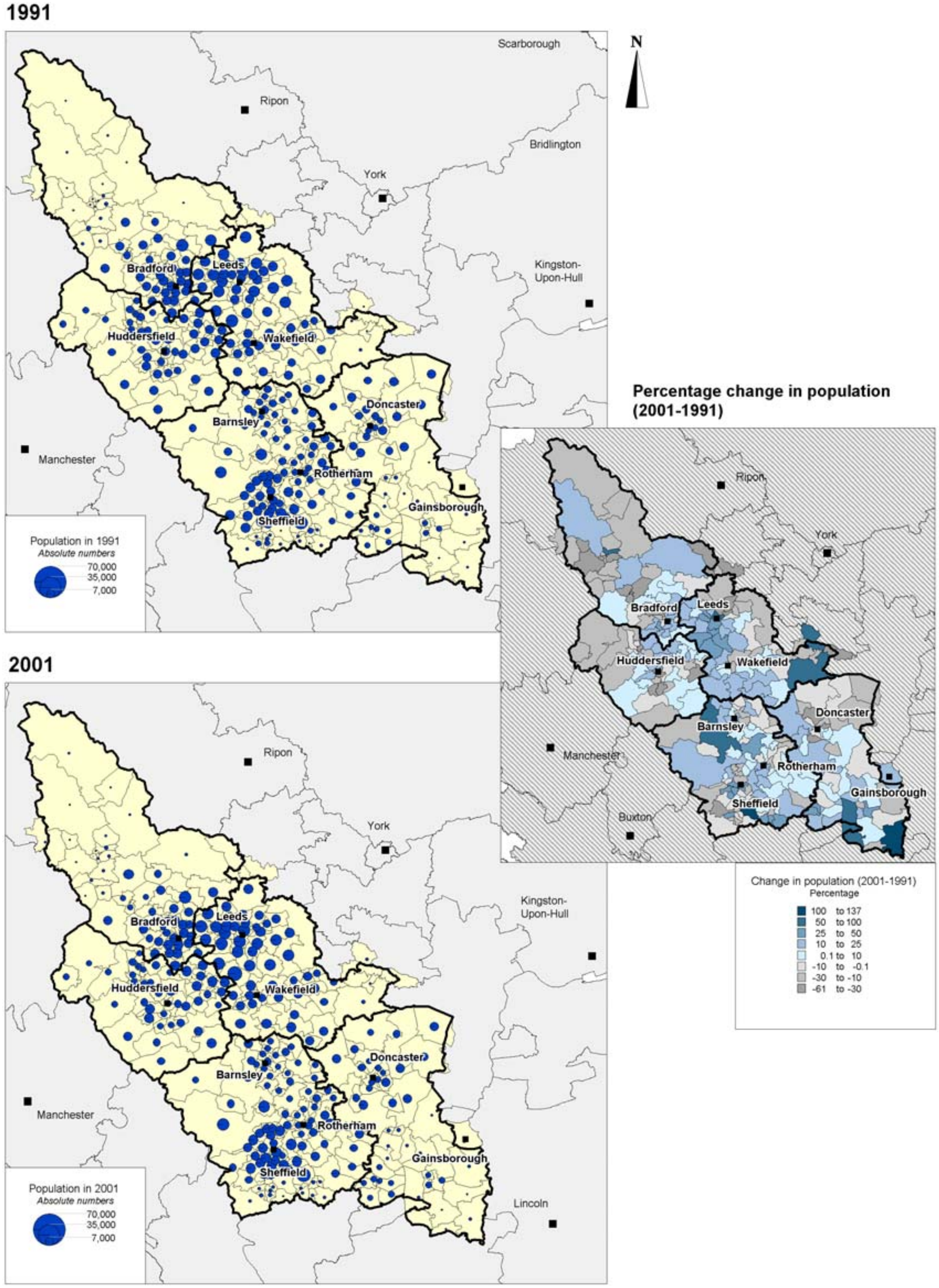

Figure 4. Population in 1991 and 2001 as predicted by the model: SimDELTA Study Area

Of the new dwellings which were introduced 510 were occupied in 1993. Figure 5 shows where the residents of these new dwellings moved from. It is immediately obvious that Barnsley residents benefit most from the new housing. Residents who previously lived in local authorities which adjoin Barnsley, in particular Rotherham occupy 27.5% of the new dwellings, whilst only 5% are occupied by households who are new to the model. Of the 340 households who previously lived in Barnsley, 130 previously lived in Wombell North.

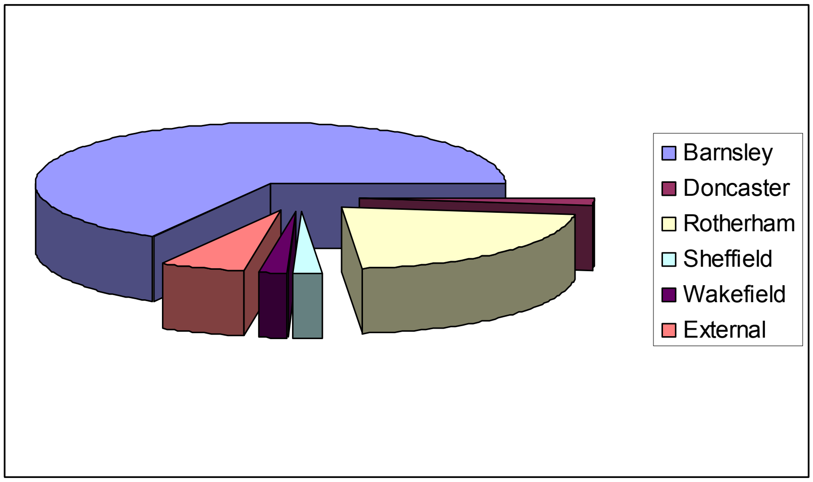

Figure 5. The addition development test: distribution of the new arrived residents in 1993.

Table 3. The additional development test: the household types of the new residents

Household type	Frequency	Percentage
Young Single	30	5.9%
Older Single	30	5.9%
Single Parent	0	0.0%
Single Retired	20	3.9%
Young Couple, no Children	50	9.8%
Older Couple, no Children	100	19.6%
Couple with Children	90	17.6%
3+ Adults, no Children	90	17.6%
3+ Adults with Children	60	11.8%

The model allocates each new dwelling a dwelling type and number of rooms, this in turn influences who will move to the area. The average number of rooms for the new dwellings was 4.8. Detached housing accounted for 21% of the new dwellings, semi-detached housing made up 50%, terraced housing made up 26% and 3% of the new dwellings were flats. Table 3 shows the household types of the new residents.

6. CONCLUSION

The main difficulty in implementing the model was the complexity of model calibration. This involves hundreds of model parameters each of which impacts on the model outputs in a way that is frequently highly correlated with that of the others, and it is very easy to get trapped in a never ending circular process fixing one problem only to find a new one pops up somewhere else. Therefore it was essential to break the calibration process into a series of logical, sequential steps, a strategy for calibration.

The modelling work described here has the important potential to contribute to understanding the consequences of planning policy and, potentially, to forecasting the impacts of possible future policies.

ACKNOWLEDGEMENTS

We are grateful to the DfT Project Officers, Russell Harris and Mo Shakharami and to members of the project Steering Group for a number of helpful discussions during the course of the project.

REFERENCES

Ballas, D. (2001) 'A spatial microsimulation approach to local labour market policy analysis', Unpublished PhD thesis, School of Geography, University of Leeds, 2001.

Clark, A.V. and Huang, Y. (2003) 'The life course and residential mobility in British housing markets' *Environmental and Planning A,* **35**, pp.323-339.

Dale, A. (1993), 'The content of the 1991 Census: change and continuity', In Dale, A., Marsh C., (eds), *The 1991 Census User's Guide*, London, HMSO, pp.16-51.

Dale, A., Fieldhouse, E., and Holdsworth, C. (2000), *Analyzing Census Microdata,* Arnold, London.

Feldman, O., Simmonds, D., Ballas, D., Clarke, G., Gibson, P., Jin, J., and Stillwell, J. (2005), 'A spatial microsimulation approach to land-use modelling', Paper presented at CUPUM conference, 29 June -1 July, 2005, London.

Mackett, R. (1990), 'MASTER Model (Micro-Analytical Simulation of Transport, Employment and Residence)', Report to the Transport and Road Research Laboratory on Contract 842/55, July 1990.

Mackett, R. (1992), 'Micro Simulation Modelling of Travel and Locational Processes: testing and further development', Report to the Transport and Road Research Laboratory, March 1992.

Mackett R. (1993), 'Micro Simulation Modelling of Travel and Locational Processes:

Results', Report to the Transport and Road Research Laboratory, May 1993.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. (1953), 'Equation of state calculations by fast computing machines', *Journal of Chemical Physics,* **21**, pp.1087-1092.

MVA (2001). *South and West Yorkshire Multi-Modal Study: Development of the Strategic Transport Model*. Report to Government Office for Yorkshire and the Humber.

MVA and DSC (2005). Troll, N., Feldman, O., Simmonds, D., Davies, G., and Tsang, F. *Balancing Areas from 2001 Census.* Final Report to DfT.

Rees, P., Martin, D., and Williamson, P. (eds) (2002), *The Census Data System*, Wiley, London.

Simmonds, D.C. and Feldman, O. (2005), 'Land-use/transport modelling with the DELTA package: progress and experience.' Paper presented at CUPUM conference, 29 June -1 July, 2005, London.

Simmonds, D.C. and Skinner, A. (2004), 'The South and West Yorkshire Strategic Model'. In Stillwell, J. and Clarke, G. (eds), *Applied GIS and Spatial Analysis,* Wiley, Chichester.